

**Analysis of Short-Term and Long-Term Variability
in Measures of Physical Function**

by

Feiyang Zheng

A thesis submitted to The Johns Hopkins University in conformity with the
requirements for the degree of Master of Science.

Baltimore, Maryland

April, 2018

© Feiyang Zheng 2018

All rights reserved

Abstract

Maintaining functional independence is crucial for older adults to have a high quality of life, as well as to reduce the risk of frailty, disability, placement into nursing homes, and debilitating diseases. It is well established that level of physical function declines with age in older adults. Hence, gerontologists have developed measurement tools to assess physical function and to monitor the changes with age. Previous investigators have conducted analysis to evaluate the sensitivity and reliability of short- and long-term measurements of the Short Physical Performance Battery (SPPB), a widely used measure of physical function. Researchers have also investigated the short-term variability of walking speed by accessing the “internal memory” between consecutive measurements. To our knowledge, no one has modeled how the correlation between measurements of physical function decays as a function of sampling time interval. Here we develop robust regression models to examine the temporal correlation of SPPB and walking speed, across large time scales, by combining short- and long-term measurements. Specifically, we developed weighted linear quantile regression models, which capture the decline in temporal correlation for the physical

ABSTRACT

function measures. We show how our method can be utilized to find the optimal sampling frequency in measures of walking speed and SPPB. We performed bootstrap resampling (500 times) to evaluate the 90% confidence interval of the optimal sampling frequency corresponding to a specified correlation. Our method can be used to support the design of observational and interventional studies with respect to physical function measurement in older people.

Primary Reader: Ravi Varadhan

Secondary Reader: Karen Bandeen-Roche

Acknowledgments

I would first like to thank my thesis advisor Prof. Varadhan for the continuous support of my research, for his patience, motivation, and immense knowledge. His guidance helped me all through the research and writing of this thesis.

Besides my advisor, I would like to thank my reader: Prof. Bandeen-Roche for her time and suggestions. She gave very detailed and helpful suggestions even though she was extremely busy as the Head of the Department of Biostatistics.

My master's degree could not have been completed without the support from my classmates in the Biostatistics Department. I appreciate their willingness to share their expertise and their time.

Finally, I must express my very profound gratitude to my parents for providing me with unfailing support and continuous encouragement throughout my years of study and through the process of researching and writing this thesis. This accomplishment would not have been possible without them. Thank you.

Contents

Abstract	ii
Acknowledgments	iv
List of Tables	viii
List of Figures	ix
1 Introduction	1
2 Methods	5
2.1 Study Population	5
2.2 Statistical Analysis	7
3 Analysis of Variability in Walking Speed	11
3.1 Assessment	11
3.2 Short-term Analysis	12
3.2.1 Exploratory Analysis	12

CONTENTS

3.3	Long-term Analysis	15
3.3.1	Exploratory Analysis	15
3.4	“Stitching Together” WHAS and WHAS substudy	17
3.4.1	Structure of Autocorrelation Function	17
3.4.2	What Is An “Optimal” Sampling Interval for Measuring Walking Speed	19
3.4.3	Alternative Way of Modeling the Data	22
3.4.4	Computing Confidence Interval for Optimal Sampling Frequency	23
3.4.5	Intraclass Correlation Coefficient	24
4	Analysis of Variability in SPPB	25
4.1	Assessment	25
4.2	Short-term Analysis	26
4.2.1	Exploratory Analysis	26
4.3	Long-term Analysis	28
4.3.1	Exploratory Analysis	28
4.4	“Stitching Together” WHAS and WHAS substudy	30
4.4.1	Structure of Autocorrelation Function	30

CONTENTS

4.4.2	What Is An “Optimal” Sampling Interval for Measuring SPPB	32
4.4.3	Alternative Way of Modeling the Data	35
4.4.4	Computing Confidence Interval for Optimal Sampling Frequency	36
4.4.5	Intraclass Correlation Coefficient	36
5	Discussion	38
	Bibliography	45
	Vita	50

List of Tables

3.1	Sample size stratified by age and disability in WHAS substudy in measures of walking speed	12
3.2	Sample size stratified by age and disability in WHAS in measures of walking speed.	15
3.3	Point estimates and 90% Confidence interval for optimal sampling frequency (weeks) for walking speed using WLQR and WLSR.	23
3.4	Intraclass correlation of WHAS and WHAS substudy in measures of walking speed.	24
4.1	Sample size stratified by age and disability in WHAS substudy in measures of SPPB.	26
4.2	Point estimates and 90% Confidence interval for optimal sampling frequency (weeks) for SPPB using WLQR and WLSR	36
4.3	Intraclass correlation of WHAS and WHAS substudy in measures of SPPB.	37

List of Figures

3.1	Walking speed (m/sec) measured each week for 24 weeks in WHAS substudy. Age groups are indicated by three different colors (red: 65-74, green: 75-84, blue: 85+). Top row shows the boxplot of walking speed at each time point; middle row depicts the actual trajectory; the bottom row shows predicted trajectories from a linear mixed-effects model	13
3.2	Walking speed (m/sec) measured every six months for 3 years in WHAS. Age groups are indicated by three different colors (red: 65-74, green: 75-84, blue: 85+). Top row shows the boxplot of walking speed at each time point; middle row depicts the actual trajectory; the bottom row shows predicted trajectories from a linear mixed-effects model	16
3.3	Correlogram for walking speed combining WHAS and WHAS substudy.	17
3.4	Correlograms for walking speed combining WHAS and WHAS substudy stratified by disability status.	18
3.5	Weighted Quantile regression for walking speed linking WHAS and WHAS substudy in terms of walking speed. Solid line: weighted linear quantile regression (WLQR); Dashed line: Weighted non-linear quantile regression	20
3.6	WLQR linking WHAS and WHAS substudy stratified by disability status in terms of walking speed.	21
3.7	WLQR and weighted least squares regression (WLSR) linking WHAS and WHAS substudy in terms of walking speed with the same weight. Solid line: WLQR; Dashed line: WLSR	23
4.1	Short physical performance battery (SPPB) measured each week for 24 weeks in WHAS substudy. Age groups are indicated by three different colors (red: 65-74, green: 75-84, blue:85+). Top row shows the boxplot of walking speed at each time point; middle row depicts the actual trajectory; the bottom row shows predicted trajectories from a linear mixed-effects model	27

LIST OF FIGURES

4.2	Short physical performance battery (SPPB) measured every 6 months for 3 years in WHAS. Age groups are indicated by three different colors (red: 65-74, green: 75-84, blue:85+). Top row shows the boxplot of walking speed at each time point; middle row depicts the actual trajectory; the bottom row shows predicted trajectories from a linear mixed-effects model	29
4.3	Correlogram combining WHAS and WHAS substudy in terms of SPPB.	30
4.4	Correlograms combining WHAS and WHAS substudy stratified by disability status in terms of SPPB.	31
4.5	Weighted quantile regression linking WHAS and WHAS substudy in terms of SPP. Solid line: weighted linear quantile regression (WLQR); Dashed line: weighted non-linear quantile regression	32
4.6	WLQR linking WHAS and WHAS substudy stratified by disability status in terms of SPPB.	34
4.7	WLQR and weighted least squares regression (WLSR) linking WHAS and WHAS substudy in terms of SPPB with the same weight. Solid line: WLQR; Dashed line: WLSR.	35

Chapter 1

Introduction

It has been well established that physical function tends to decline with age among older adults. Previous studies have highlighted that the decline of physical performance limits the ability of older adults to participate in family and community activities [1]. Lack of physical activity also reflects serious problems in public health [2]. Lower levels of physical activity are associated with increased risk of frailty, disability, institutionalization, dementia and Alzheimers disease [3–6]. Therefore, it is crucial for older adults to sustain physical independence to reduce the risk of diseases and to have a high quality of life. Moreover, monitoring the change of physical functions will help researchers to develop the strategy in terms of identifying the risk factors for reduced physical function. There are several assessment tools for evaluating the physical performance of older adults. In this thesis, we focus on two of the most widely accepted measures: walking speed and short physical performance battery (SPPB).

CHAPTER 1. INTRODUCTION

Walking speed is an easily and quickly assessed performance measure requiring minimal training of the participants [7]. Researchers considered walking as a good indicator of the “slowing down process” which accompanies aging. Studies have also suggested that walking speed tends to decline more discernibly in older adults between ages 65-70 years, in both females and males [8]. In addition, walking speed plays an essential role in predicting of multiple health-related outcomes including hospitalization, falls and even mortality [6, 9]. For example, based on the pooled analysis of nine US cohorts, every 0.1 m/s increase in walking speed results in 12% decrease in hazard to death [10, 11]. Hence, walking speed is an important measure of overall physical function [12], and is most widely used in gerontologic research studies [7].

The Short Physical Performance Battery (SPPB) is an objective measurement of lower extremity physical performance status [13]. SPPB consists of 3 components: standing balance, walking speed, and chair stand test. Each of the tests is rescaled based on a standard score system and aggregated to a summary SPPB score ranging from 0 to 12 [13]. Specifically, 0 means poor performance and 12 means high performance. In several studies, SPPB has been shown to be a powerful predictor of adverse outcomes [14]. For example, one study indicated that non-disabled people with low SPPB scores could be up to 5 times more likely to have disability in the activities of daily living or mobility-related disability than those with high summary scores, four years later [15]. Besides, compared with self-reported measures of physical function,

CHAPTER 1. INTRODUCTION

objective measures such as walking speed and SPPB, offer obvious advantages since they are more reproducible and sensitive to change [16].

Repeated, longitudinal measurements are taken to monitor the changes in physical performance as the increase of people’s age. Therefore, it is important to understand the measurement properties of these assessment tools. The reliability of walking speed and SPPB has been evaluated by Ostir [14], who assessed the reliability and sensitivity of both long-term and short-term measurements of SPPB. Ferrucci et al. suggested to investigate the short-term variability of walking speed by accessing the “internal memory” between consecutive measurements [17]. In addition, the National Institute on Aging (NIA) has called for research to study interventions to slow age-related declines physical function in older adults. To design such study, it is critical to choose an appropriate sampling frequency for monitoring changes in function. Too frequent measurements can result in redundant information that can be prohibitively expensive; and too infrequent sampling can result in failure to track important changes in function. Currently, there is no method for estimating “optimal” sampling frequency.

Our work builds on the papers by Ferrucci et al and Ostir et al by modeling the temporal correlation in the measures of walking speed and SPPB, across short- and long-term measurement time scales [14, 17]. Specifically, we develop a novel approach that “stitches” together the short- and long-term data, and derive robust linear regression models that capture the decay in temporal correlation for the physical

CHAPTER 1. INTRODUCTION

function measures. Our model equations can be used to estimate “optimal” sampling frequency for measuring walking speed and SPPB in future studies of physical function in older adults. Our approach of combining short-term and long-term longitudinal data can also be employed for modeling other physical and cognitive function measures in older adults.

Chapter 2

Methods

2.1 Study Population

The data used for in this thesis came from two studies. The long-term data were from the Womens Health and Aging Study (WHAS) while the short-term data was also from the same WHAS cohort, but from a subset of participants who participated in the WHAS weekly substudy [18]. WHAS was a community-based study conducted by The Johns Hopkins Medical Institutions and sponsored by the Laboratory of Epidemiology, Demography, and Biometry of the National Institute on Aging. In this study, 5,316 out of 32,538 subjects living in 12 contiguous ZIP code areas in Baltimore, MD were selected from the Health Care Financing Administrations Medicare enrollment file for screening. The recruiting criteria were women 65 years of age or older with moderate to severe disability [19].

CHAPTER 2. METHODS

4,137 subjects who were able to participate in the screening process were screened according to cognitive functioning and disability status in order to identify the one third most disabled older women living in the community. The disability assessment contained 15 activities that were categorized into 4 domains: (1) mobility and exercise tolerance (walking 1/4 mile, walking up 10 steps without resting, getting in and out of bed or chairs, doing heavy housework); (2) upper extremity function (raising arms up over head, using fingers to grasp or handle, lifting and carrying 10 lb); (3) basic self-care tasks (bathing, dressing, eating, using the toilet); and (4) higher functioning tasks (using telephone, doing light housework, preparing meals, shopping for personal items) [17]. Women were asked to report whether they had any difficulty completing each of these activity by themselves without help. 1,409 women reported difficulty completing activities in 2 or more domains and were also scored greater than 17 on the Mini-Mental State Examination, a global test of cognitive function [20]. Among these 1002 subjects agreed to participate in the study. They received further assessment including physical function evaluation and health examination every 6 months for 3 years in their homes.

In the substudy, 105 subjects in the second and third replicates in WHAS were selected consecutively until each of 9 cells according to age (65-74, 75-84, and 85+) and disability status (difficulty in two, three, or four domains as described above) had approximately equal numbers of subjects. Participants were assessed objectively in physical function and asked about health condition weekly for 24 weeks in their

homes. In this thesis, “WHAS” indicates the long-term study and “WHAS substudy” indicates the short-term study.

2.2 Statistical Analysis

Participants’ demographic characteristics were summarized using $\text{mean} \pm \text{standard deviation}$ and range for continuous variables, while count and percentage were calculated for categorical variables. Following Ferrucci et al’s analysis, box plots and spaghetti plots were displayed to present the population trend and the individuals’ trajectory of walking speed and SPPB stratified disability domain. Mixed effect models with random effect and slope were fitted for participants with distinct disability status to observe the individual change of both walking speed and SPPB [17]. Models with fixed effect were also fitted to illustrate the population trends stratified by disability status

We computed the Pearson correlation “r” between pairs of measurements in a time series separated by time interval of different lengths. This is known as the correlogram [21]. We calculated the correlogram for each time lag in both WHAS substudy and WHAS. For example, suppose we have K visits: For the lag of 1, we computed the correlation between $y(t_i)$ and $y(t_{i+1}), i = 1 \dots K - 1$; for the lag of 2, we computed the correlation between $y(t_i)$ and $y(t_{i+2}), i = 1 \dots K - 2$; and so forth. When calculating the correlation in each pair of measurements, we only included

CHAPTER 2. METHODS

people without missing values. The median correlation was calculated for pairs of measurements within each lag, i.e. for lag $k, k = 1, 2, \dots, K - 1$, we computed the median of $K - k$ pairwise correlations. The times were equally spaced in each of the studies. In the WHAS substudy there were 24 visits, each 1 week apart, giving us a total of 23 lags. In WHAS, there were 7 visits, each 6 months apart, giving us a total of 6 lags. Because the time lags cover a large range of values, we used the logarithmic scale for better visualization. Correlogram was also constructed for subjects in different disability groups.

The median correlation was modeled by quantile linear regression as a function of the log of the time lag. There are several reasons to use linear quantile regression. First, since the number of pairs at time lag is not large, we believed that the median is a more robust estimator of central measure of correlation than mean. Second, there are outliers in our data and quantile regression is more robust to outliers. Third, inference based on quantile regression makes fewer assumptions about underlying data distribution than least squares regression [22, 23]. Thus, our approach is robust both for using median correlations and by modeling these using quantile regression.

Considering the assumption that $\text{var}(r) \sim 1/N$, the study sample size was used as the weights. For walking speed, the weight for WHAS substudy was calculated as $102/(102+1002)$ and the weight for WHAS was $1002/((102+1002))$. Similarly, when we estimated the correlation in terms of SPPB, the weight on WHAS substudy was calculated as $105/(105+1002)$ and the weight on WHAS was $1002/((105+1002))$. For

CHAPTER 2. METHODS

walking speed, the information of 3 subjects was lost, so we had a smaller sample size. The median correlation was also estimated by stratifying subjects according to disability domain.

In order to assess the adequacy of the linearity assumption, we added a restricted cubic splines with 3 knots to the quantile regression, which allows for the time trend to be nonlinear. Furthermore, we also fitted a weighted least squares regression to demonstrate the robustness of quantile regression to outliers.

Optimal sampling frequency was estimated from the fitted quantile linear regression model for correlation as the time lag corresponding to the desired correlation between successive measurements. Nonparametric Bootstrap was used to obtain a 90% confidence interval for optimal sampling frequency in different correlations. For both the short and long term studies, we resampled new datasets that have the same sample size of the original ones with replacement. For each new dataset, we computed the Pearson correlation at different time lags and fitted a weighted quantile linear regression, as well as a weighted least squares regression. We then estimated the optimal sampling frequencies corresponding to correlations of 0.85, 0.8, 0.75 and 0.7. This process was repeated 500 times for both walking speed and SPPB. We sorted the computed optimal measurement frequency and defined the 0.05 percentage as the lower bound and the 0.95 percentage as the upper bound of the 90% confidence interval.

Following Ostir's analysis [14], we also computed the one-way intraclass correla-

CHAPTER 2. METHODS

tion, which is defined as the ratio of between-person variance to the sum of between-person variance and within-person variance. This is a standard metric of reliability. Within-person variance refers to the change or fluctuation within an individual over time, and between-person variation refers to the difference among people in their average physical performance [24]. ICC may be classified into 4 categories which are poor(<0.40), fair (0.40-0.59), good (0.60-0.74) and excellent(0.75-1.00) [25].

Chapter 3

Analysis of Variability in Walking Speed

3.1 Assessment

Walking speed was assessed by timing participants walking 3 or 4 meters across a small room at their usual pace. The test was repeated twice, and the better performance was recorded. The unit of walking speed is meter per second. 0 indicates that the woman was unable to perform the test [17].

3.2 Short-term Analysis

3.2.1 Exploratory Analysis

Information on age and disability was collected for participants at baseline. The total sample size was 105. However, 3 subjects missed information on walking speed, so the analytic sample size was 102 subjects in WHAS substudy. Of the 102 subjects, the average age of the subjects was 78.2 ± 8 years (Range: 66-95). There were 33 participants who had disability in 2 domains, 35 who had disability in 3 domains, and 34 who reported disability in 4 domains. Subjects were stratified into 9 groups based on their age and disability domain shown in table 3.1. The participants were predominately white(60%). The mean walking speed was 0.95 ± 0.38 m/sec (Range: 0.061-3.1).

Table 3.1: Sample size stratified by age and disability in WHAS substudy in measures of walking speed

Age Group	Disability in 2 domains	Disability in 3 domains	Disability in 4 domains
65-74 yrs	10	14	10
75-84 yrs	11	14	11
85+ yrs	12	7	13

Figure 3.1 displays plots of walking speed measured for each participant in WHAS substudy [17]. The three columns represent study population with disability in 2 do-

CHAPTER 3. ANALYSIS OF VARIABILITY IN WALKING SPEED

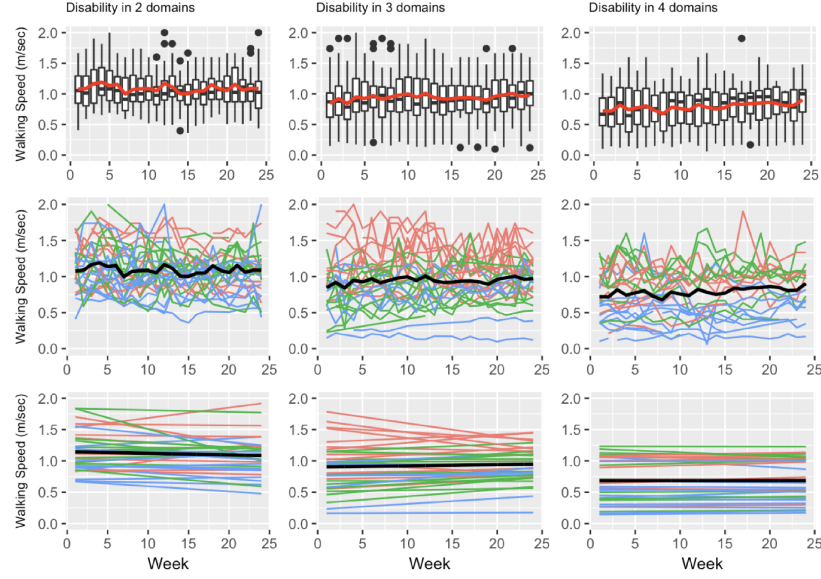


Figure 3.1: Walking speed (m/sec) measured each week for 24 weeks in WHAS sub-study. Age groups are indicated by three different colors (red: 65-74, green: 75-84, blue: 85+). Top row shows the boxplot of walking speed at each time point; middle row depicts the actual trajectory; the bottom row shows predicted trajectories from a linear mixed-effects model

mains, 3 domains, and 4 domains respectively. The first row is the box plot providing visual summaries of distributions including the third quantile, median, the first quantile, and outliers. Red line shows average walking speed (m/sec). Based on the top plots, in general, average walking speed is relatively stable in the first group. For other two groups, a slight increase can be observed and the growth rate rises with the severity of disability. Moreover, the first column also has the highest walking speed and the lowest variability.

The second row summarizes the individuals' walking speed trajectories for each age and disability group. Measurements that belong to the same person are connected by a line and age groups are indicated by three different colors (red: 65-74, green:

CHAPTER 3. ANALYSIS OF VARIABILITY IN WALKING SPEED

75-84, blue:85+). The black line represents the average walking speed (m/sec) of each week. The 3 age groups evenly distribute between 0.3 m/sec to 2 m/sec with a similar average in the first column. In the second column, people in the youngest age group have remarkably high values. In the third column, it is evident that the oldest people have the worst performance while the distribution of other two groups overlaps. We can observe high fluctuations in the individual trajectories of people with the youngest and intermediate age with all levels of disability. However, individual trend gradually becomes more stable as the increase of disability domains for people older than 85 years.

In the bottom plots, weekly walking speed is predicted by mixed effect models for each individual assuming random intercept and slope. The black line is the prediction of walking speed with fixed effect for the entire population of each subgroup. The slope of the line illustrates how quickly walking speed changes. The majority of individuals with the disability in 4 domains reports a roughly constant walking speed across time. However, in other two groups, there are more individual differences. The trends stay roughly constant at the population level in all groups generally.

3.3 Long-term Analysis

3.3.1 Exploratory Analysis

Information on age and disability was collected for participants at baseline. Of the 1002 subjects in WHAS, the average age of the participants was 78.3 ± 8.1 years (range: 65-101 years). 372 of these participants had disability in 2 domains, 267 had disability in 3 domains, and 363 reported disability in 4 domains. Subjects were stratified into 9 cells based on their age and disability groups shown in table 3.2. Regarding race, the study population only contained white and black, which accounted for 71% and 29% respectively. The mean walking speed was 0.57 ± 0.38 m/sec (Range: 0-2.67).

Table 3.2: Sample size stratified by age and disability in WHAS in measures of walking speed.

Age Group	Disability in 2 domains	Disability in 3 domains	Disability in 4 domains
65-74 yrs	175	101	112
75-84 yrs	109	90	112
85+ yrs	88	76	139

Figure 3.2 displays plots of walking speed measured for each participant in WHAS. In the top plots, average walking speed is highly stable without fluctuation across time in all disability groups. However, the variability enlarges along with time. Comparing the three groups, there is no evident difference between walking speed for the least

CHAPTER 3. ANALYSIS OF VARIABILITY IN WALKING SPEED

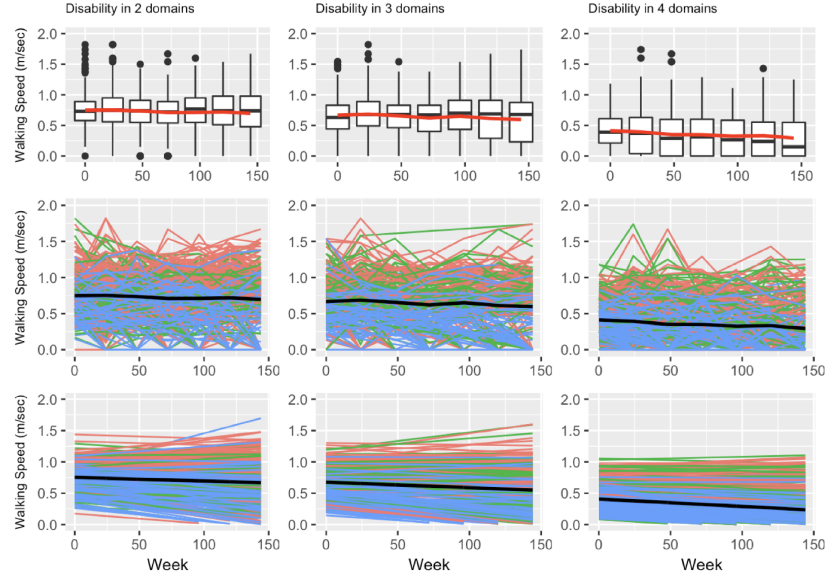


Figure 3.2: Walking speed (m/sec) measured every six months for 3 years in WHAS. Age groups are indicated by three different colors (red: 65-74, green: 75-84, blue: 85+). Top row shows the boxplot of walking speed at each time point; middle row depicts the actual trajectory; the bottom row shows predicted trajectories from a linear mixed-effects model

disabled group and the intermediate disabled group. Walking speed of people in the highest disability status have a dramatic drop with “floor effect”. The variability also becomes greater and greater moving from the left column to the right column. The comparison demonstrates a similar impact of low and intermediate disability status and a significant negative effect of high disability status on walking speed.

For plots in the second row, older people are located on the bottom while other two age groups mix together. The distribution suggests that after a certain age, the walking speed of women declines rapidly.

In the bottom plots, the trend suggests a slight decrease across time for all three groups in long follow-up time. Significant growth of a few individuals can be seen

CHAPTER 3. ANALYSIS OF VARIABILITY IN WALKING SPEED

in the first and second columns while the trajectories are more homogeneous for the individuals in the third column.

The same plots were plotted for 25 samples randomly selected from each disability group for the purpose of visualization in the appendix.

3.4 “Stitching Together” WHAS and WHAS substudy

3.4.1 Structure of Autocorrelation Function

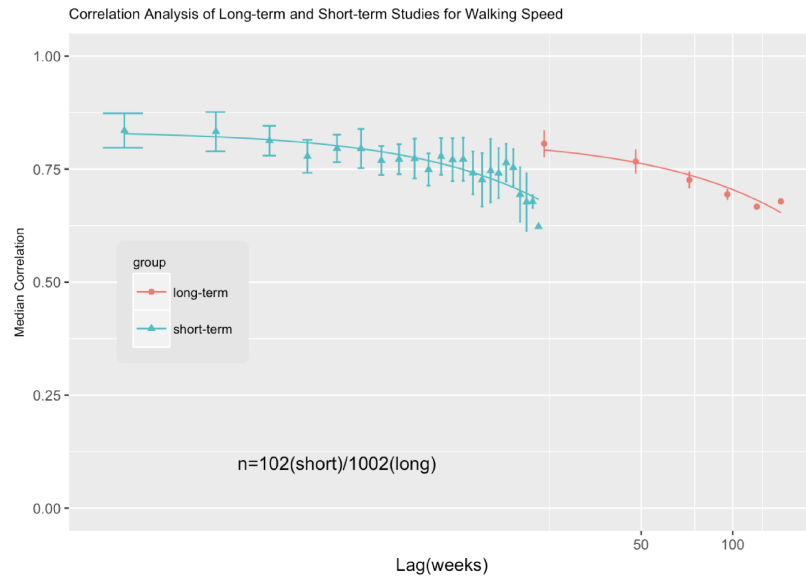


Figure 3.3: Correlogram for walking speed combining WHAS and WHAS substudy.

Figure 3.3 shows the correlogram of walking speed with error bars between mea-

CHAPTER 3. ANALYSIS OF VARIABILITY IN WALKING SPEED

measurements as a function of time lag combining WHAS and WHAS substudy. The blue color presents WHAS substudy and the red color presents WHAS. The y axis indicates median correlation and the x axis indicates time lags from 1 to 144 in log scale. Considering WHAS and WHAS substudy separately, the correlation between measurements is high at small time lag and decays as the increase of time lag. However, when we regarding the two data sets as a whole, there is a disparity between the two studies. The plot roughly demonstrates the change of correlation across time linking the two studies together. Because the sample size of the two studies is not comparable, this plot can be misleading especially for points after 20 lags on the substudy due to outliers. As a consequence, we fitted quantile regression and put weight based on the sample size proportion as shown in the next subsection.

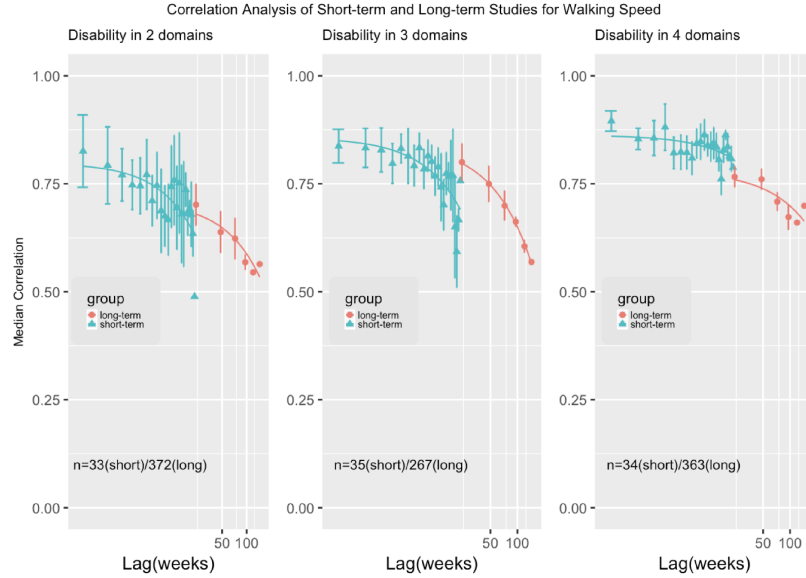


Figure 3.4: Correlograms for walking speed combining WHAS and WHAS substudy stratified by disability status.

CHAPTER 3. ANALYSIS OF VARIABILITY IN WALKING SPEED

Figure 3.4 reveals the correlation between pairs of measurements for participants stratified by disability domain as a function of time lag. Similarly, the correlation between measurements is high at small time lag and decays as the increase of time lag if we consider WHAS and WHAS substudy separately. Error bars in the first two groups show great variability of correlation in short-term analysis. In addition, we can observe substantially low values at around 20 weeks in the middle plots. Therefore, much of the disparity in the last section seems to be driven by the 3 domain group. Comparing the 3 plots, we may roughly conclude that for people with greater disability, the measurement frequency does not need to be as high to maintain the correlation at the same level as before. Said differently, we can have less frequent sampling for people with more disability since they appear to exhibit more stability.

3.4.2 What Is An “Optimal” Sampling Interval for Measuring Walking Speed

The solid line in figure 3.5 demonstrates the weighted linear quantile regression (WLQR) to predict the median correlation at a specific time lag. We can observe some outliers at around 20 weeks, which is consistent with the findings from the last subsection. The dashed line indicates the weighted non-linear quantile regression with restricted cubic spline added to the time trend. Clearly, the two lines highly overlap showing the linearity assumption is not violated.

CHAPTER 3. ANALYSIS OF VARIABILITY IN WALKING SPEED

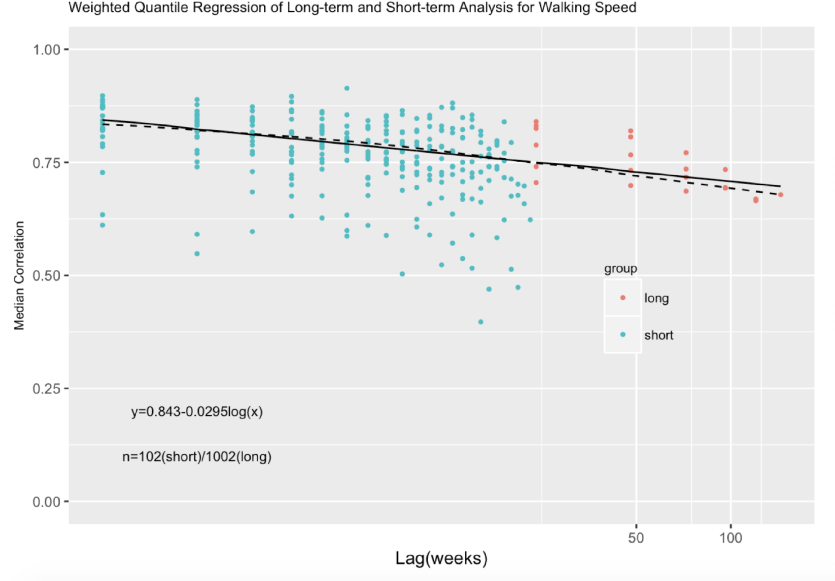


Figure 3.5: Weighted Quantile regression for walking speed linking WHAS and WHAS substudy in terms of walking speed. Solid line: weighted linear quantile regression (WLQR); Dashed line: Weighted non-linear quantile regression

The equation based on our model is

$$y = 0.843 - 0.0295 \log(x) \quad (3.1)$$

as shown in the plot. In the above, y is the median correlation between consecutive measurements while x is the time lag (in weeks) between two measurements. This equation can be used to determine the optimal measurement frequency given a time lag in week. The way that we define an “optimal” frequency depends on the objective of a study. For example, if we want to monitor the decline of walking speed in a study, a very high correlation can result in too frequent and redundant sampling, wasting time and resources. In this case, suppose a correlation of 0.75 is optimal, the analysis suggests to maintain a measurement frequency to be around $\exp[(0.843 -$

CHAPTER 3. ANALYSIS OF VARIABILITY IN WALKING SPEED

$0.75)/(0.0295)] \simeq 24$ weeks (around 6 months) according to the equation. Conversely, we can estimate the correlation between measurements when a study has been designed. If a study measures walking speed of subjects every 8 week, a correlation of 0.78 can be expected.

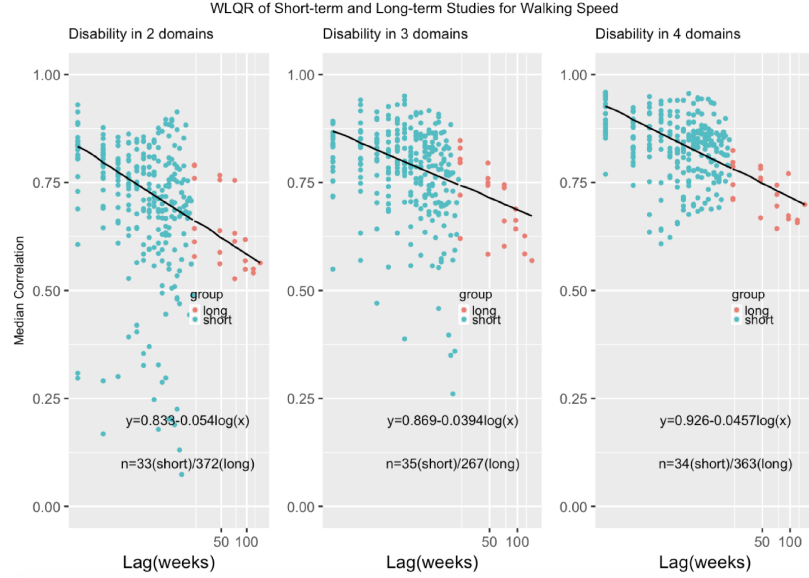


Figure 3.6: WLQR linking WHAS and WHAS substudy stratified by disability status in terms of walking speed.

Figure 3.6 contains the WLQR to predict the correlation at a specific time lag for the 3 disability groups. Each subplot indicates a different disability group. Similarly, subjects belong to distinct disability groups could use the corresponding functions for higher accuracy. The equations for participants in the 3 disability groups are

$$y = 0.833 - 0.054 \log(x) \quad (3.2)$$

$$y = 0.869 - 0.0394 \log(x) \quad (3.3)$$

$$y = 0.926 - 0.0457 \log(x) \quad (3.4)$$

CHAPTER 3. ANALYSIS OF VARIABILITY IN WALKING SPEED

Equation 3.2 is used for people with disability in 2 domains. Equations 3.3 and equation 3.4 are used for people with 3 and 4 domains respectively. These equation indicates an increase of starting correlation with the severity of disability. Using the example from the previous subsection, if we want to maintain a correlation of 0.75, the measurement frequencies for the 3 groups are $\exp[(0.833-0.75)/(0.054)] \simeq 4$ weeks (1 month), $\exp[(0.869-0.75)/(0.0394)] \simeq 20$ weeks (5 months) and $\exp[(0.926-0.75)/(0.0457)] \simeq 47$ weeks (12 months). Clearly, as the increase of the number of disability domains, the measurement frequency can decrease to sustain the same correlation.

3.4.3 Alternative Way of Modeling the Data

An alternative way of modeling the data is weighted least squares regression (WLSR) with the same weight shown in figure 3.7. The solid line presents the WLQR while the dashed line indicates the WLSR. WLQR fits line for median correlation while WLSR fits line for mean correlation. It is obvious that the mean correlation predicted using WLSR is lower than the median correlation predicted by WLQR when the time lag is smaller than 23 weeks. The two lines intercept at 24 weeks and highly overlap. Mean correlation in the short-term analysis has a greater variability and more outliers, which drives the WLSR downwards. Therefore, because WLQR is more robust to outliers, it has a better performance to model walking speed.

CHAPTER 3. ANALYSIS OF VARIABILITY IN WALKING SPEED

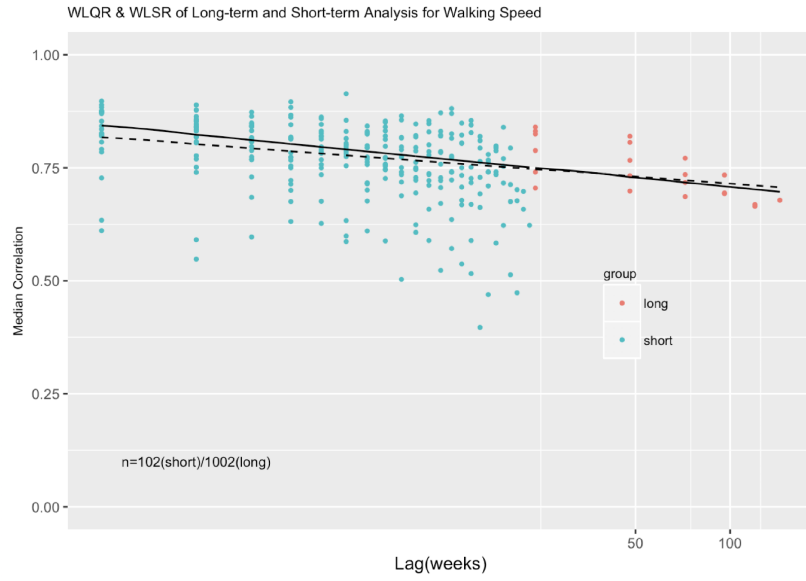


Figure 3.7: WLQR and weighted least squares regression (WLSR) linking WHAS and WHAS substudy in terms of walking speed with the same weight. Solid line: WLQR; Dashed line: WLSR

3.4.4 Computing Confidence Interval for Optimal Sampling Frequency

Table 3.3: Point estimates and 90% Confidence interval for optimal sampling frequency (weeks) for walking speed using WLQR and WLSR.

Desired correlation	WLQR	WLSR
0.85	<1 (<1,4)	<1 (<1,<1)
0.8	4 (<1,14)	2 (<1,15)
0.75	23 (7,53)	21 (1,63)
0.7	127 (57,454)	195 (64,2266)

CHAPTER 3. ANALYSIS OF VARIABILITY IN WALKING SPEED

Table 3.3 illustrates the 90% confidence intervals and the original point estimations for optimal sampling frequency in weeks in measures of walking speed using WLQR and WLSR. The 90% confidence interval was estimated by bootstrapping 500 times. As we can see, interval estimated using WLQR is significantly narrower and more stable. When the correlation is greater than 0.7, confidence intervals obtained using the two models are similar. However, the confidence interval becomes extremely large when the correlation equals to 0.7. In this case, WLQR has a much narrower confidence interval than WLSR.

3.4.5 Intraclass Correlation Coefficient

Table 3.4: Intraclass correlation of WHAS and WHAS substudy in measures of walking speed.

Study	ICC
WHAS	0.72
Substudy	0.77

Table 3.4 indicates the intraclass correlation coefficient for walking speed measurements in both WHAS and WHAS substudy. The results assesses the reliability of the two studies. The ICC in WHAS substudy is 0.77, which is indicated as “excellent” reliability. The ICC in WHAS is 0.72 and can be evaluated as “good” reliability.

Chapter 4

Analysis of Variability in SPPB

4.1 Assessment

SPPB includes: walking speed (better of two timed 3 or 4 meter walks across a small room), one repeated chair stand test (Time for 5 chair stands), and balance stands (the ability to stand in a tandem position, semi-tandem position, or side-by-side position for 10 seconds) [14]. A 5-level summary score from 0 to 4 of each performance test was created after calculation based on specific criteria. 0 indicates that the subjects are unable to perform the test while 4 indicates the best performance. The three measurements were added to create a SPPB score with value from 0 to 12. If more than one score of the three components were missing, the SPPB score of this subject was set to be missing. However, if only one score was missing, the SPPB score was defined as the sum of other two non-missing measurements plus

the average of the two non-missing scores [19].

4.2 Short-term Analysis

4.2.1 Exploratory Analysis

Information on age and disability was collected for participants at baseline. There were no missing values, so the analytic sample size was 105. Of the 105 subjects in the WHAS substudy, the average age of the subjects was 78.8 ± 7.9 years (Range: 66-95 years). 33 of these participants had disability in 2 domains, 37 had disability in 3 domains, and 35 reported disability in 4 domains. Subjects were stratified into 9 groups based on their age and disability domain shown in table 4.1. The mean SPPB was 7.22 ± 3.39 (Range: 0-12).

Table 4.1: Sample size stratified by age and disability in WHAS substudy in measures of SPPB.

Age Group	Disability in 2 domains	Disability in 3 domains	Disability in 4 domains
65-74 yrs	10	14	10
75-84 yrs	11	14	11
85+ yrs	12	9	14

Figure 4.1 displays plots of SPPB measured for each participant in WHAS sub-study. The three columns represent study population with disability in 2 domains, 3

CHAPTER 4. ANALYSIS OF VARIABILITY IN SPPB

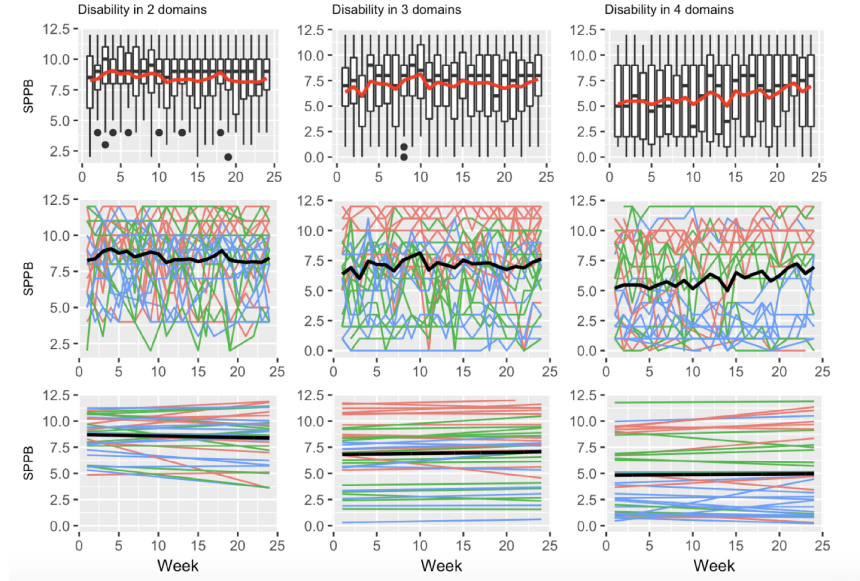


Figure 4.1: Short physical performance battery (SPPB) measured each week for 24 weeks in WHAS substudy. Age groups are indicated by three different colors (red: 65-74, green: 75-84, blue: 85+). Top row shows the boxplot of walking speed at each time point; middle row depicts the actual trajectory; the bottom row shows predicted trajectories from a linear mixed-effects model

domains, and 4 domains respectively. The first row is the box plot providing visual summaries of distributions including the third quantile, median, the first quantile, and outliers. Red line shows average SPPB. According to the first row, average SPPB scores are relatively stable over time for subjects in both the intermediate and the lower disability groups. Obviously, people with the highest disability status have the lowest average score at all time points and there is a significant improvement over time. Moreover, the interquartile range rises with the increase of disability domain, suggesting a greater variability for people with high disability status.

The second row summarizes the individuals' SPPB score trajectories for each age and disability group. Measurements that belong to the same person are connected

CHAPTER 4. ANALYSIS OF VARIABILITY IN SPPB

by a line and age groups are indicated by three different colors (red: 65-74, green: 75-84, blue:85+). The black line represents the average SPPB of each week. Inside each plot of the second row, younger people are located on the top while older people are located on the bottom seen clearly in the second and third plots. This disparity becomes more and more visible with the severity of disability.

In the bottom plots, weekly SPPB score is predicted by mixed effect models for each individual assuming random intercept and slope. The black line is the prediction of SPPB with mixed effect for the entire population of each subgroup. The slope of the line illustrates how quickly SPPB changes. In the bottom plots, the trend suggests a slightly increase across time in the intermediate group with a positive slope of most individuals. There are more crossovers between lines in other two groups suggesting greater individual difference. The change of general population appear to be close to 0 for all the subgroups indicating no change in short period of time.

4.3 Long-term Analysis

4.3.1 Exploratory Analysis

The sample size and the demographic information are the same as walking speed measurement shown in section 3.3.1. The mean SPPB was 5.43 ± 3.71 (Range:0-12).

Figure 4.2 shows plots of SPPB scores in WHAS. Within each plot, significant increase of interquartile range is visualized across time from the first visit to the last

CHAPTER 4. ANALYSIS OF VARIABILITY IN SPPB

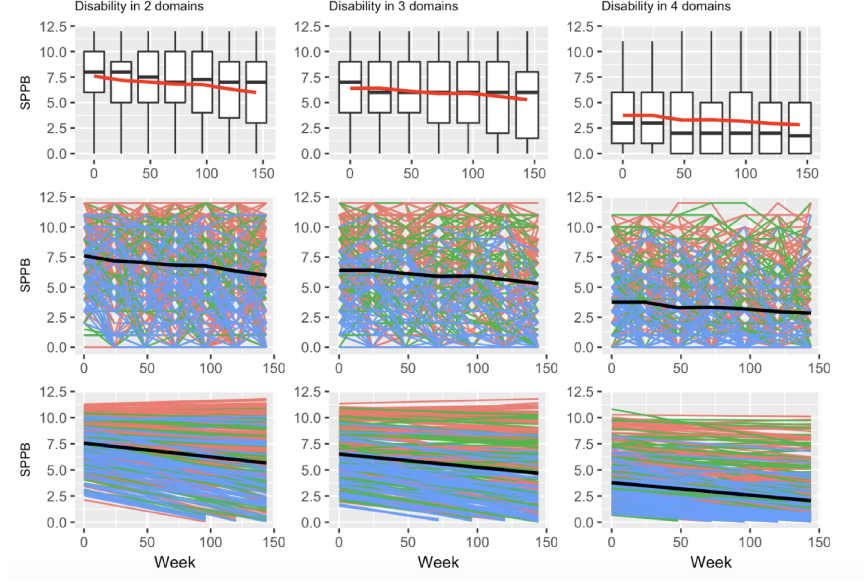


Figure 4.2: Short physical performance battery (SPPB) measured every 6 months for 3 years in WHAS. Age groups are indicated by three different colors (red: 65-74, green: 75-84, blue:85+). Top row shows the boxplot of walking speed at each time point; middle row depicts the actual trajectory; the bottom row shows predicted trajectories from a linear mixed-effects model

visit. “Floor effect” occurs in the group with 4 disability domains due to the low SPPB score. In general, a gradual decline of SPPB scores could be observed moving from the left column to the right. The drop is highly substantial in the most disabled group comparing with other two.

In the middle plot, age does not significantly affect SPPB score for subjects with the lowest disability status. However, with the rise of disability severity, oldest people generally move to the bottom.

There is a remarkable agreement for all the subgroups with a decay in SPPB scores across time on a similar speed in population level. Therefore, in long follow-up time, the SPPB score tends to decay for all subjects in the study.

CHAPTER 4. ANALYSIS OF VARIABILITY IN SPPB

Same plots were plotted for 25 samples randomly selected from each disability group for the purpose of visualization in the appendix.

4.4 “Stitching Together” WHAS and WHAS substudy

4.4.1 Structure of Autocorrelation Function

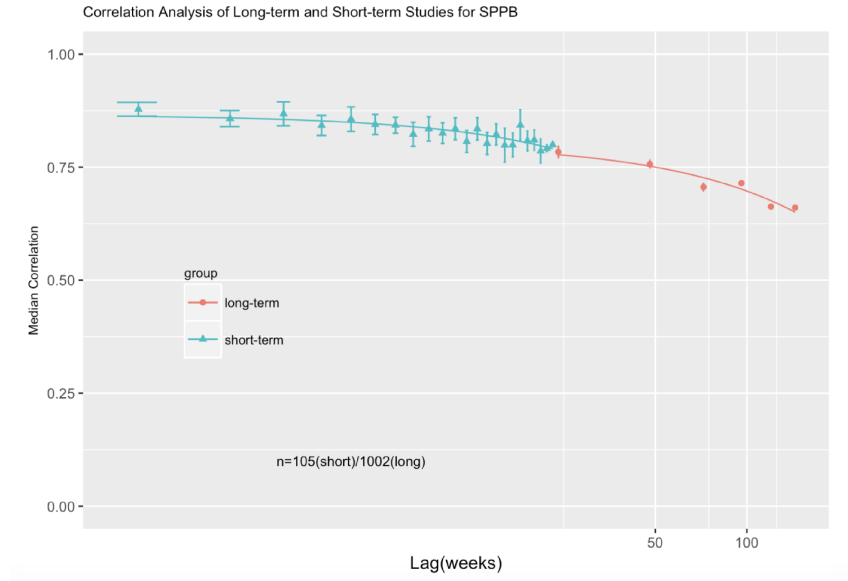


Figure 4.3: Correlogram combining WHAS and WHAS substudy in terms of SPPB.

Figure 4.3 shows the correlogram of SPPB score with error bars between measurements as a function of time lag combining WHAS and WHAS substudy. The blue color presents WHAS substudy and the red color presents WHAS. The y axis

CHAPTER 4. ANALYSIS OF VARIABILITY IN SPPB

indicates median correlation and the x axis indicates time lags from 1 to 144 in log scale. The correlation between measurements is high at small time lag and decays as the increase of time lag. The median correlations between the two studies are well connected at around 0.78. In this case, the maximum median correlation is 0.86 with one-week time separation while the minimum median correlation equals to 0.66 when the time lag is 144. It is also noticeable that short-term analysis has a larger error bar indicating a greater variability.

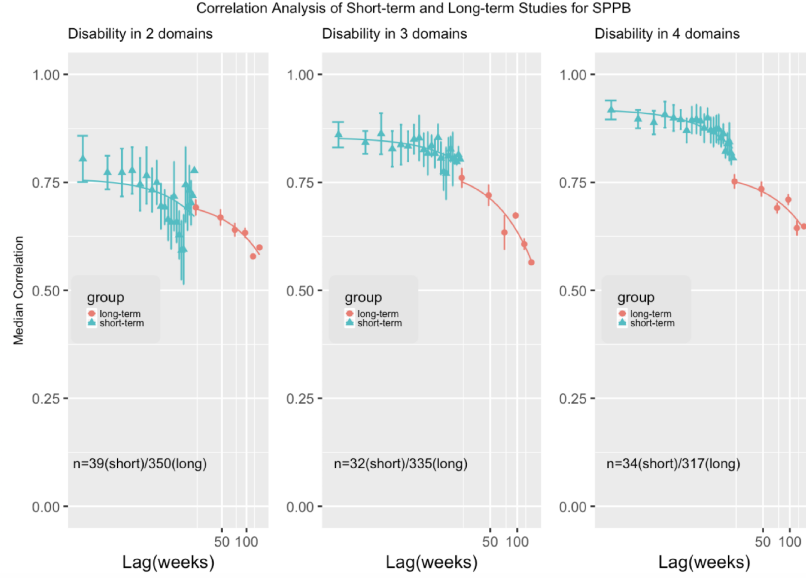


Figure 4.4: Correlograms combining WHAS and WHAS substudy stratified by disability status in terms of SPPB.

Figure 4.4 reveals the correlation between pairs of measurements for participants stratified by disability domain as a function of time lag. Clearly, the median correlation declines as the increase of time lag in both WHAS substudy and WHAS in all groups. When we regard the two studies as a whole, the result is consistent. In other

CHAPTER 4. ANALYSIS OF VARIABILITY IN SPPB

words, the median correlation and the decline rate grow with the disability severity demonstrating positive associations. We may also conclude that as the increase of the disability severity, the measurement frequency can be reduced to maintain the same correlation.

4.4.2 What Is An “Optimal” Sampling Interval for Measuring SPPB



Figure 4.5: Weighted quantile regression linking WHAS and WHAS substudy in terms of SPP. Solid line: weighted linear quantile regression (WLQR); Dashed line: weighted non-linear quantile regression

The solid line in figure 4.5 demonstrates the weighted linear quantile regression (WLQR) to predict the median correlation at a specific time lag. The dashed line indicates the weighted non-linear quantile regression with restricted cubic spline added

CHAPTER 4. ANALYSIS OF VARIABILITY IN SPPB

to the time trend. There seems to be some non-linearity for SPPB, but linear assumption is still reasonable..

The equation based on our model is

$$y = 0.91 - 0.042 \log(x) \quad (4.1)$$

as shown in the plot. In the equation above, y is the median correlation between pairs of measurements while x is the time lag (in weeks) between two measurements. This equation can be used to determine the optimal measurement frequency for the purpose of study design. Using the same example as walking speed, in order to keep a correlation to be 0.75 to catch the decline of SPPB scores, we are suggested to maintain a measurement frequency to be at least $\exp[(0.91-0.75)/(0.042)] \simeq 45$ weeks (around 11 months). This function can also be applied to predict the median correlation at a specific time lag. If the SPPB of subjects are reported every 8 week, we would expect a correlation to be around 0.82.

Figure 4.6 contains the WLQR to predict the correlation at a specific time lag for the 3 disability groups. Each subplot indicates a different disability group. Similarly, subjects belong to distinct disability groups could use the corresponding functions for higher accuracy. The equations for participants in the 3 disability groups are

$$y = 0.813 - 0.0409 \log(x) \quad (4.2)$$

$$y = 0.925 - 0.0551 \log(x) \quad (4.3)$$

$$y = 0.994 - 0.068 \log(x) \quad (4.4)$$

CHAPTER 4. ANALYSIS OF VARIABILITY IN SPPB

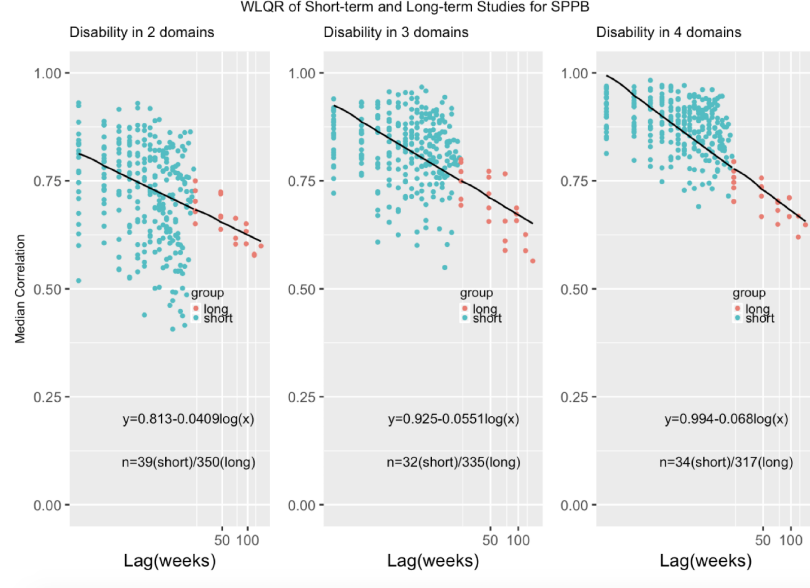


Figure 4.6: WLQR linking WHAS and WHAS substudy stratified by disability status in terms of SPPB.

Equation 4.2 is used for people with disability in 2 domains. Equations 4.3 and equation 4.4 are used for people with 3 and 4 domains respectively. These equations indicate an increase of starting correlation as the severer of disability. Using the example from the previous subsection, if we want to maintain a correlation of 0.75, the measurement frequencies for the 3 groups are estimated to be $\exp[(0.813-0.75)/(0.0409)] \simeq 4.7$ weeks (1 month), $\exp[(0.925-0.75)/(0.0551)] \simeq 24$ weeks (6 months) and $\exp[(0.994-0.75)/(0.068)] \simeq 36$ weeks (9 months). Clearly, as the increase of disability domain, the measurement frequency can be reduced to maintain the same correlation as before.

CHAPTER 4. ANALYSIS OF VARIABILITY IN SPPB

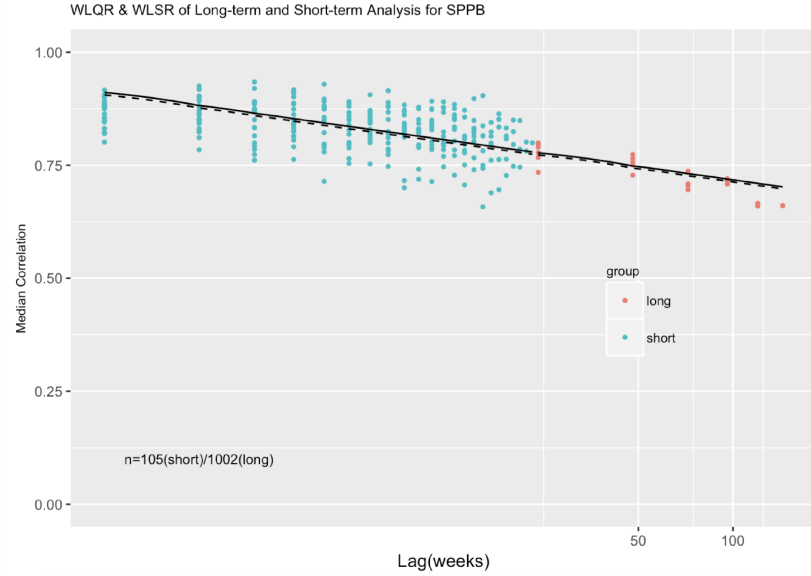


Figure 4.7: WLQR and weighted least squares regression (WLSR) linking WHAS and WHAS substudy in terms of SPPB with the same weight. Solid line: WLQR; Dashed line: WLSR.

4.4.3 Alternative Way of Modeling the Data

Besides WLQR, we also fitted a weighted least squares linear regression (WLSR) shown in figure 4.7 to compare. The solid line indicates WLQR and the dashed line indicates the WLSR. WLQR fits the median correlation, while WLSR fits the mean correlation. It is well known that least-squares fitting is sensitive to outliers in the data. Here the two lines are quite close to each other since there are no outliers in the data.

4.4.4 Computing Confidence Interval for Optimal Sampling Frequency

Table 4.2: Point estimates and 90% Confidence interval for optimal sampling frequency (weeks) for SPPB using WLQR and WLSR

Desired correlation	WLQR	WLSR
0.85	4 (<1,8)	4 (<1,8)
0.8	14 (4,23)	12 (2,22)
0.75	45 (18,78)	41 (13,78)
0.7	148 (71,426)	136 (74,408)

Table 4.2 illustrates the 90% confidence intervals and the original point estimations for optimal sampling frequency in weeks in measures of SPPB using WLQR and WLSR. The 90% confidence interval was estimated by bootstrapping 500 times. The two methods show similar performance, which is consistent with figure 4.7 in the last subsection. Thus, when there are no substantial outliers, both of the two methods are acceptable.

4.4.5 Intraclass Correlation Coefficient

Table 4.3 shows the intraclass correlation coefficient for the summary performance score in both WHAS and WHAS substudy. The ICC in WHAS substudy is 0.84,

CHAPTER 4. ANALYSIS OF VARIABILITY IN SPPB

Table 4.3: Intraclass correlation of WHAS and WHAS substudy in measures of SPPB.

Study	ICC
WHAS	0.71
Substudy	0.84

which is indicated as “excellent” reliability. The ICC in WHAS is 0.71 and can be evaluated as “good” reliability.

Chapter 5

Discussion

Monitoring the change of physical functions can help researchers identify the risk factors for reduced physical function. There are several assessment tools for evaluating the physical performance of older adults. In this thesis, we focused on two of the most widely used measures of physical function: walking speed and short physical performance battery (SPPB).

We investigated the variability of walking speed and SPPB for women 65 years or older with disability in 2 to 4 domains using WHAS (long-term) and WHAS substudy (short-term). Our work extended papers by Ferrucci et al and Ostir et al by modeling the temporal correlation in the measures of walking speed and SPPB, across short- and long-term measurement time scales [14, 17]. Specifically, we developed a novel approach that “stitched” together the short- and long-term data, and derived robust linear regression models that capture the decay in temporal correlation for the physical

CHAPTER 5. DISCUSSION

function measures. We used the model equations to estimate “optimal” sampling frequency for measuring walking speed and SPPB in future studies of physical function in older adults.

We keep patterns of the functional trajectory of women in different age and disability groups in terms of walking speed and SPPB. On average, people with disability in 4 domains showed a greater decline in physical functions compared with those in 2 and 3 domains. Larger variability and “floor effect” can also be observed in this group. This pattern is extremely compelling in long-term analysis indicating a possibility that 3 domains might be a threshold for good physical performance. In other words, for people with disability in 4 domains, their life quality in terms of physical functions may be heavily impacted due to disability.

Another interesting and important observation is that both the mean SPPB performance and mean walking speed were substantially better in the short-term sub-study compared to the long-term main study. The mean walking speed in the short-term and long-term studies are 0.95 m/s and 0.57 m/s, respectively; and the mean SPPB scores in the short-term and long-term studies are 7.22 and 5.43, respectively. This suggests that the people who enrolled into the rigorous, weekly follow-up for 24 weeks, are a healthier subset of the main study. It is possible that the autocorrelation structure of the healthier subset of women may be different than the autocorrelation structure of others. This is a form of selection bias. Our analyses that stitched together the short- and long-term measurements did not take this into account. It may

CHAPTER 5. DISCUSSION

not be entirely possible to eliminate this bias from a recruitment perspective, since enrolling a representative population might result in a larger proportion of missing data. Analytic strategies to address this bias need to be developed. It would be interesting to look at other characteristics of this select group such as coexisting chronic conditions, as well as frailty status, as these are potential factors which could explain the selection [5]. Thus, our current approach to stitching together the short- and long-term measurements would need to be extended to include these factors which partially account for the selection bias.

Although the physical function remains stable in the short-term, it decreases in long-term follow-up. People with poor performances initially will continue to remain low functioning or even deteriorate faster with time. Interestingly, the decrease in function for the long-term analysis is more dramatic for SPPB than for walking speed. Hence, we may conclude that SPPB is more sensitive to age-related changes.

The change in performance of an individual can be captured using repeated measurements. Thus, correlogram in figure 3.3 and figure 4.3 constructed to compute the autocorrelation implies the degree of individual heterogeneity across time. Obviously, a high sampling frequency can perfectly catch each tiny change. However, unnecessary measurement is a waste of resources and may also introduce more missing values. Here, in both studies, the correlation remained quite high: Even a lag of 144 weeks yielded a correlation above 0.66, which suggests that a very high sampling frequency can be redundant in measures of walking speed and SPPB. Furthermore, after strat-

CHAPTER 5. DISCUSSION

ifying participants according to disability status, we found that the correlations were different across the disability groups, with the highest correlation in the most disabled group. Thus, designing the measurement frequency based on disability status may be considered an appropriate way to explain the heterogeneity in physical function.

We modeled the autocorrelation as a function of the log of the time lag in order to estimate the measurement frequency. Outliers are clearly seen in the figure 3.3 when the time lag equals approximately 20 weeks, with significantly low correlations because of small sample size. To avoid the potentially large influence of outliers, we employed a linear quantile regression approach. We also weighted the short-term and long-term studies proportional to their sample sizes to roughly reflect the variance of median correlation estimated for each lag. Based on equation 3.1, in order to maintain a correlation of 0.75, walking speed is suggested to be measured every 6 months. For SPPB, according to equation 4.1, 11 months is enough to capture the changes. The results are consistent with our previous observation that SPPB is more sensitive to changes. Moreover, weighted linear quantile regression (WLQR) was fitted for each disability group separately. Based on our estimation, in order to maintain the correlation in the same level, we would allow a lower measurement frequency for people with higher disability status. An alternative way of modeling the data is weighted least squares regression (WLSR). When modeling the autocorrelation of SPPB, the WLSR performs as well as the WLQR. However, when there are outliers, as is the case for walking speed, the WLQR seems to perform better (compare the

CHAPTER 5. DISCUSSION

point estimates and the width of the confidence intervals).

We also make an interesting remark about the confidence intervals for the optimal sampling frequency. We computed it using bootstrap. The optimal sampling frequency estimate is calculated by inverting the best-fit linear regressions in equations (3.1) and (4.1). This estimator is a ratio of two estimators. Estimating the confidence interval for a ratio of normally distributed estimators is known as the Fieller-Creasy problem [26]. The estimator for the optimal sampling frequency does not have a finite expectation, when the confidence interval of the denominator includes zero. This also explains the large variance of the estimator for optimal sampling frequency. It is also unclear as to whether the bootstrap approach is even applicable due to infinite expectation. Development of a theoretically rigorous approach to estimating the confidence interval for this estimator is a topic for further study.

Another modeling approach is to transform the autocorrelation and then perform the regression. Fisher's z-transformation, $\log \frac{1+r}{1-r}$ or even a logit transformation, $\log \frac{r}{1-r}$ are good options. The transformed model might provide a better fit to the data, and furthermore, it might yield narrower confidence intervals.

The reliability and sensitivity to change of SPPB and walking speed measurement were explored using the intraclass correlation coefficient (ICC) for both WHAS and WHAS substudy. Specifically, the ICC values for SPPB and walking speed in the main study of WHAS are similar. However, in the substudy, the ICC of SPPB is 0.07 higher than that of walking speed. SPPB measurement consists of 3 domains,

CHAPTER 5. DISCUSSION

including walking speed, and hence it stands to reason that it is more stable.

Our methodology for combining the short-term and long-term measurements in order to estimate an “optimal” sampling frequency is novel. It may be applied to evaluate other physical and cognitive functions. It should also be mentioned that there could be different objective functions to optimize. Here we chose the simple criterion of achieving a desired autocorrelation. Another possible approach is to come up with the smallest sampling frequency that would enable us to identify a change of a given magnitude, beyond measurement error, in walking speed or SPPB. This would entail consideration of additional information including the nature of trajectories (e.g., linear or quadratic decline) and inter-person variability in the rate of decline. This is an interesting area for further development.

A main limitation of our methodology is that it requires both short-term and long-term measurements spanning a wide range of sampling time scale. We seldom have such data available. Therefore, we suggest that future studies build this type of data collection into their design in order to assess the measurement properties of novel assessments. One such example is the assessment of low-grade proinflammation in older adults, which is measured using cytokines like interleukin-6 (IL6) and C-reactive protein (CRP). The optimal sampling frequency for these measures is not known. Therefore, it would be ideal to design a short-term substudy, nested within the long-term cohort study. Then, our approach can be used to determine an appropriate sampling frequency.

CHAPTER 5. DISCUSSION

Another limitation of our study is missing data. Both of the two studies contain intermittent missingness and drop out, especially in short-term study since subjects were measured more frequently. In our analysis, we only kept people without missing in each of two measurements at different time lag to calculate correlation. In future study, we may want to do some imputation to replace missing data with imputed estimates.

In conclusion, our study examined the change of physical functions of older women in short-term and long-term measures of walking speed and SPPB. Evidence was presented to confirm that all of the measurements are reliable enough to monitor the change. The characteristic of correlation as a function of time lag was fully described by correlogram linking the two studies together. We modeled this correlogram using a quantile linear regression approach. Most importantly, we constructed equations, which enable us to determine the optimal measurement frequency for achieving a desired autocorrelation for assessing physical function for the purpose of scientific study design.

Bibliography

- [1] K. Avlund, R. Lund, B. E. Holstein, and P. Due, “Social relations as determinant of onset of disability in aging,” *Archives of Gerontology and Geriatrics*, vol. 38, pp. 85–99, 2004.
- [2] J. M. Guralnik, L. P. Fried, and M. E. Salive, “Disability as a public health outcome in the aging population,” *Annual Review of Public Healths*, vol. 17, pp. 25–46, 1996.
- [3] L. Wang, E. B. Larson, J. D. Bowen, and G. van Belle, “Performance-based physical function and future dementia in older people,” *Arch Intern Med.*, vol. 166, pp. 1115–1120, 2006.
- [4] J. Gawel, D. Vengrow, J. Collins, S. Brown, A. Buchanan, and C. Cook, “The short physical performance battery as a predictor for long term disability or institutionalization in the community dwelling elderly population aged 65 years or older.” *Phys Ther Rev*, vol. 17, p. 3744, 2012.

BIBLIOGRAPHY

- [5] L. P. Fried and et al., “Frailty in older adults: evidence for a phenotype.” *J Gerontol A Biol Sci Med Sci.*, vol. 56, pp. 146–56., 2001.
- [6] S. A. Scherer, W. R. Hiatt, and J. G. Regensteiner, “Lack of relationship between gait parameters and physical function in peripheral arterial disease.” *J Vasc Surg.*, vol. 44, pp. 782–8, 2006.
- [7] J. E. Graham, G. V. Ostir, S. R. Fisher, and K. J. Ottenbacher, “Assessing walking speed in clinical research: a systematic review.” *J Eval Clin Pract.*, vol. 14, pp. 552–62, 2008.
- [8] J. E. Himann, D. A. Cunningham, P. A. Rechnitzer, and D. H. Paterson, “Age-related changes in speed of walking.” *Med Sci Sports Exerc.*, vol. 20, pp. 161–6, 1988.
- [9] S. Studenski, S. Perera, D. Wallace, and et al., “Physical performance measures in the clinical setting.” *J Am Geriatr Soc.*, vol. 51, p. 314322, 2003.
- [10] S. Studenski, S. Perera, K. Patel, C. Rosano, K. Faulkner, and et al., “Gait speed and survival in older adults.” *JAMA*, vol. 305, pp. 50–8, 2011.
- [11] A. Elbaz, S. Sabia, E. Brunner, and et al., “Association of walking speed in late midlife with mortality: results from the whitehall ii cohort study,” *Age (Dordr)*, vol. 35, p. 943952, 2013.

BIBLIOGRAPHY

- [12] S. Studenski, S. Perera, K. Patel, and et al., “Gait speed and survival in older adults.” *JAMA*, vol. 305, p. 5058, 2011.
- [13] J. M. Guralnik, E. M. Simonsick, L. Ferrucci, R. J. Glynn, L. F. Berkman, D. G. Blazer, P. A. Scherr, and R. B. Wallace, “A short physical performance battery assessing lower extremity function: association with self-reported disability and prediction of mortality and nursing home admission.” *J Gerontol.*, vol. 49, pp. 85–94, 1994.
- [14] G. V. Ostir, S. Volpato, L. P. Fried, P. Chaves, and J. M. Guralnik, “Reliability and sensitivity to change assessed for a summary measure of lower body function,” *J Clin Epidemiol.*, vol. 55, pp. 916–21, 2002.
- [15] J. M. Guralnik, L. Ferrucci, E. M. Simonsick, M. E. Salive, and R. B. Wallace, “Lower-extremity function in persons over the age of 70 years as a predictor of subsequent disability.” *N Engl J Med.*, vol. 332, pp. 556–61, 1995.
- [16] J. M. Guralnik, L. G. Branch, S. R. Cummings, and J. D. Curb, “Physical performance measures in aging research,” *Journal of Gerontology*, vol. 44, pp. 141–146, 1989.
- [17] L. Ferrucci, J. M. Guralnik, M. E. Salive, L. P. Fried, K. Bandeen-Roche, D. B. Brock, E. M. Simonsick, M. C. Corti, and S. L. Zeger, “Effect of age and severity of disability on short-term variation in walking speed: the women’s health and aging study,” *J Clin Epidemiol.*, vol. 49, pp. 1089–96, 1996.

BIBLIOGRAPHY

- [18] J. M. Guralnik, L. P. Fried, E. M. Simonsick, J. D. Kasper, and M. E. Lafferty, *The Womens Health and Aging Study. Health and social characteristics of older women with disability*. National Institute on Aging, 1995.
- [19] G. V. Ostir, S. Volpato, L. P. Fried, P. Chaves, and J. M. Guralnik, “Reliability and sensitivity to change assessed for a summary measure of lower body function: results from the woman’s health and aging study.” *J Clin Epidemiol.*, vol. 55, pp. 916–21, 2002.
- [20] I. McDowell, B. Kristjansson, G. B. Hill, and R. Hbert, “Community screening for dementia: the mini mental state exam (mmse) and modified mini-mental state exam (3ms) compared.” *J Clin Epidemiol.*, vol. 50, pp. 377–83, 1997.
- [21] P. Diggle, P. Heagerty, K. Y. Liang, and S. Zeger, *Analysis of Longitudinal Data*. Oxford University Press, 2002.
- [22] O. J. Onyedikachi, “Robustness of quantile regression to outliers.” *American Journal of Applied Mathematics and Statistics*, vol. 3, pp. 86–88, 2015.
- [23] R. Koenker, *Quantile Regression*. Cambridge University Press, 2005.
- [24] D. K. Mroczek, A. Spiro, and D. M. Almeida, “Between- and within-person variation in affect and personality over days and years: How basic and applied approaches can inform one another,” *Ageing International*, vol. 28, pp. 28–260, 2003.

BIBLIOGRAPHY

- [25] D. V. Cicchetti, “Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology.” *Psychological Assessment*, vol. 6, pp. 284–290, 1994.
- [26] E. C. Fieller, “Some problems in interval estimation.” *Journal of the Royal Statistical Society, Series B.*, vol. 16, p. 175185, 1954.

Vita

Feiyang Zheng was born on December 21st, 1992 in China. She received the Bachelor of Science degree in Biological Science from University of California Irvine in 2016 and enrolled in the Biostatistics ScM program at Johns Hopkins University. During the time studying at Johns Hopkins University, she also worked as a biostatistician at the Center on Aging and Health, focusing on the physical function of older adults.